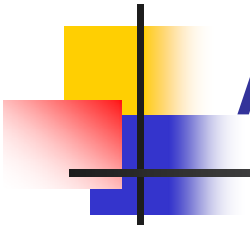
Peer-to-Peer Science Data Environment (P2PSDE)

Code 587 / Matt Holland

The Information Systems Division

Goals, Objectives, Benefits

- **What?**
  - Developing a P2P networking environment featuring real-time-discoverable science data services, searchably organized by topic-area.
- **Why?**
  - When "information" is needed, the details of where and how it is stored are usually irrelevant.
  - Scientists often spend time in data acquisition that could be otherwise used for analysis work.



Approach

- **JXTA** is a set of open, generalized P2P protocols that allow any connected devices on the network to communicate and collaborate as peers. (http://www.jxta.org)
  - JXTA protocols independent of programming language; we use the Java (J2SE) binding.
  - Our classes are designed as reusable modules; forming any applications needed for creating, maintaining and accessing the SDE.

Approach (continued)

- The SDE is a simple hierarchy:
  - single "parent" group (P2PSDE root-group)
  - unlimited number of "child" (sub-)groups.
- Peers discover what groups are available to "join" (connect). The idea is that each group represents an interest-area (related data):
  - Within groups, peers browse for data, serve data (provide value-added services), or act as "super-peer"--where super-peers support discovery.

Approach (continued)

- In general, groups have a set of services offered by peers within that group:
  - Abstract away the details of exactly *who* is serving the content, and *how* it is stored.
  - Users search for needed "information", specifying *any* relevant format for result.
  - A variety of portal (browsing) applications may be constructed from our basic Java modules.
  - Simplest portal would simply download content discovered from appropriate group's services.

Results, Status, Next Steps

- For now, peers serve locally stored files:
  - search for served files by *filename* only
  - *as stored on server*; no value-added services
- Ultimately, many services are planned:
  - search queries may describe *content* of files
  - files served in *user-specified* storage format
  - **Killer app.--IDL file-access dialogs**:
    - Analysis tools (via IDL's JavaBridge) retrieve needed science data discovered on remote peer network.
    - Scientist users may rapidly acquire remote data in exactly the form their analysis tool requires.